Улучшение автоматической стилизации изображений с помощью преобразующей нейросети

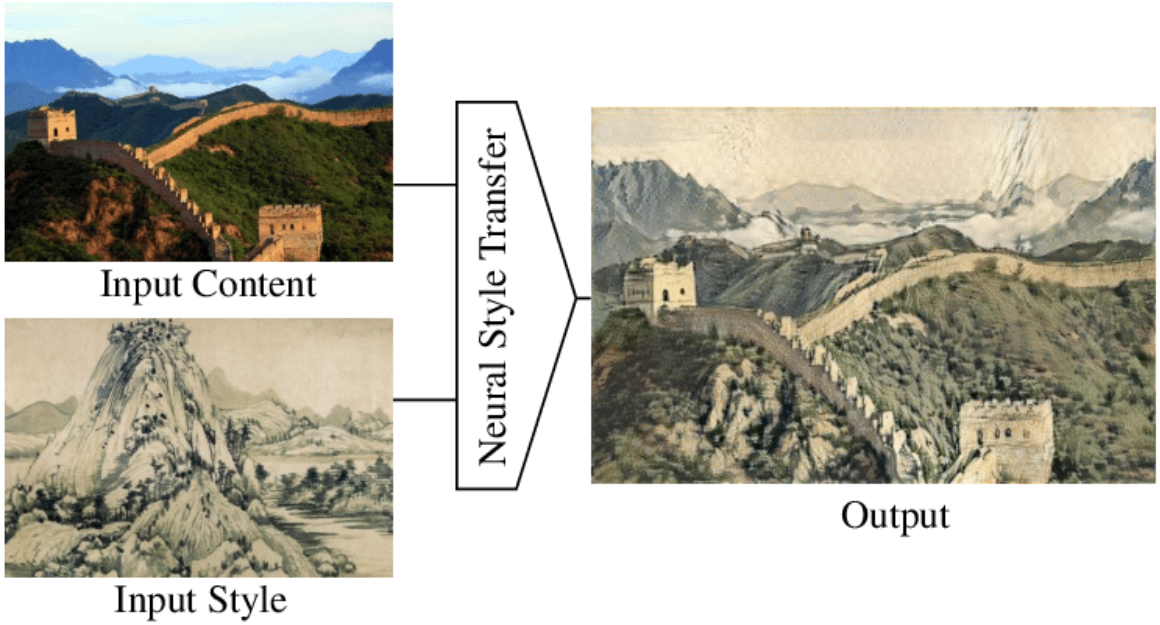
Дипломная работа

Елистратов Семен, 417
Научный руководитель:
к.ф-м.н., доцент
Китов Виктор Владимирович

ММП, ВМК

2022

Постановка задачи автоматического переноса стиля



- Дано:
  - Изображение контента: I_c
  - Изображение стиля: I_s
- Результат: стилизованное изображение контента: I_r

Оптимизационный подход

- Gatys et al. "A Neural Algorithm of Artistic Style"(2015) [1]
- Оптимизация в пространстве изображений:

$$I_r = \arg\min_x \mathcal{L}_{content}(I_c, x) + \lambda \mathcal{L}_{style}(I_s, x)$$

  - Функция потерь контента

$$\mathcal{L}_{content}(I_c, I_r) = \sum_{i \in L_\mathcal{C}} \frac{1}{n_i} \|\phi_i(I_r) - \phi_i(I_c)\|_F^2,$$

  - Функция потерь стиля

$$\mathcal{L}_{style}(I_s, I_r) = \sum_{i \in L_\mathcal{S}} \frac{1}{n_i} \|\mathcal{G}(\phi_i(I_r)) - \mathcal{G}(\phi_i(I_s))\|_F^2,$$

  где $L_\mathcal{S}$ and $L_\mathcal{C}$ - множество соответствущих слоев сети VGG, $\phi_i(\cdot)$ - внутреннее представление на i-ом слое, n_i - соответствующее число элементов тензора

Оптимизационный метод

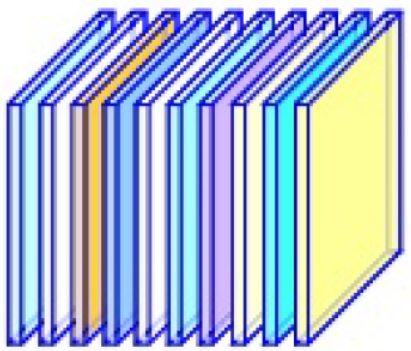


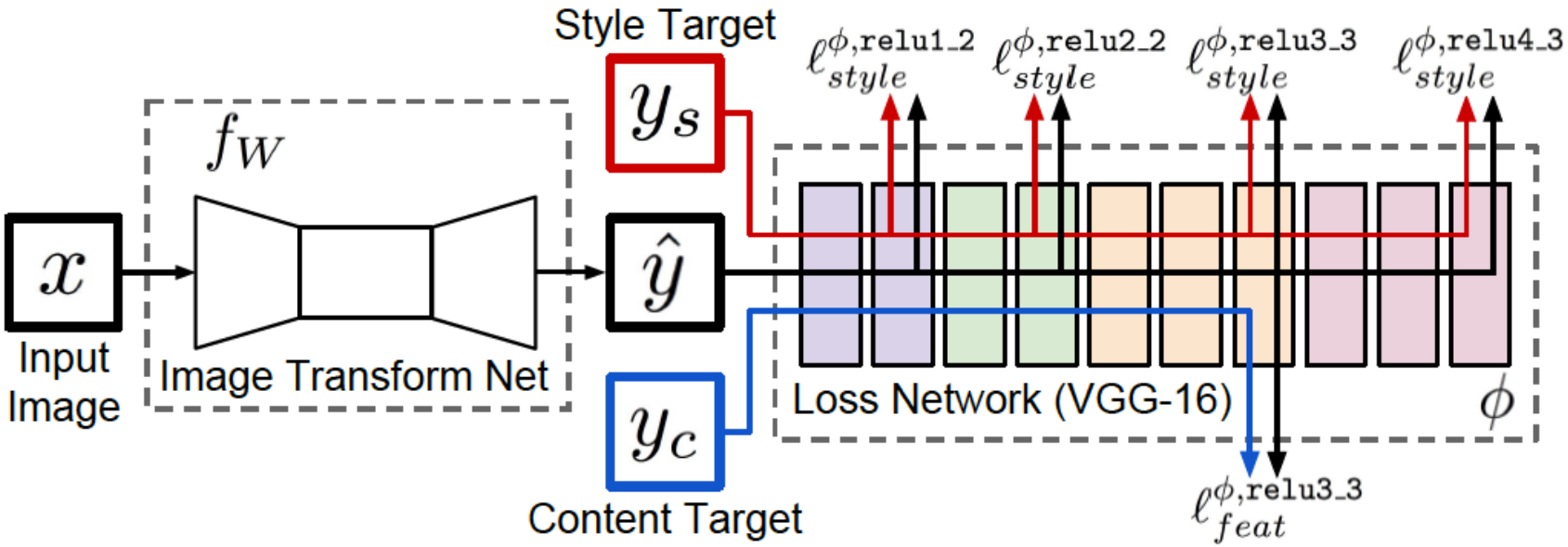
Рис.: Внутреннее представление тензора в сети VGG.

- В качестве пространственных статистик - матрицы Грама
- Элемент cc' матрицы Грама:

$$\mathcal{G}(\phi_i(x))_{cc'} = \frac{1}{C_i H_i W_i} \sum_{h=1}^{H_i} \sum_{w=1}^{W_i} \phi_i(x)_{hwc} \phi_i(x)_{hwc'}$$

- Матрица Грама - матрица ковариаций между каналами

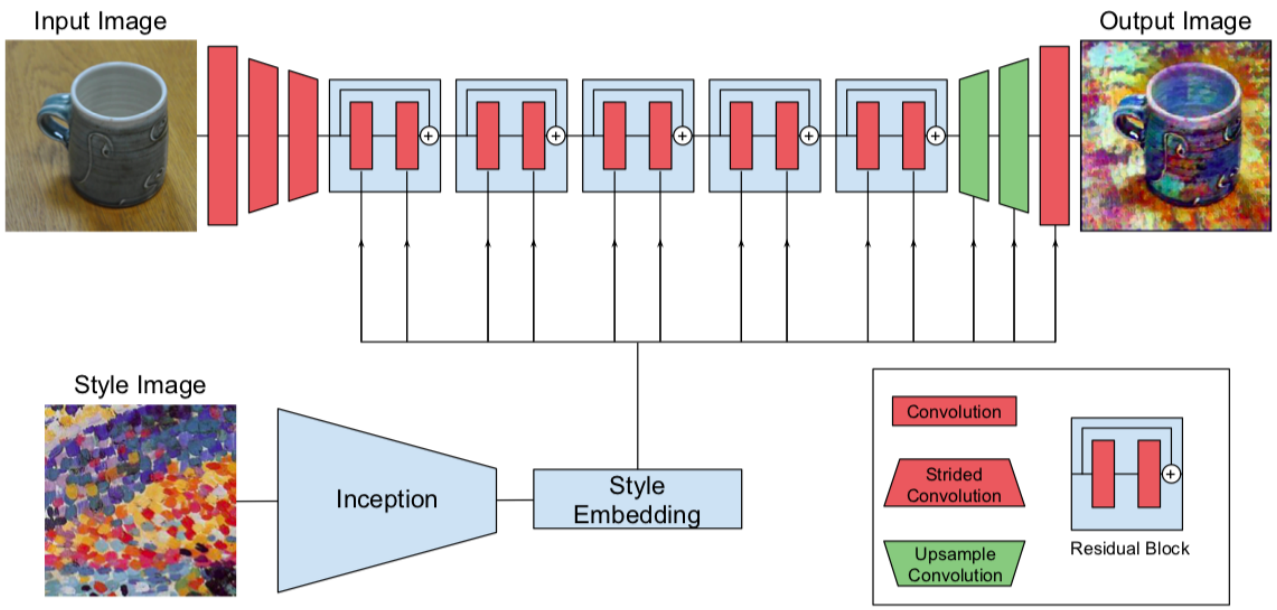
Базовый алгоритм



- Dumoulin et al. "A learned representation for artistic style."(2016) [2]
- Условная пространственная нормализация:

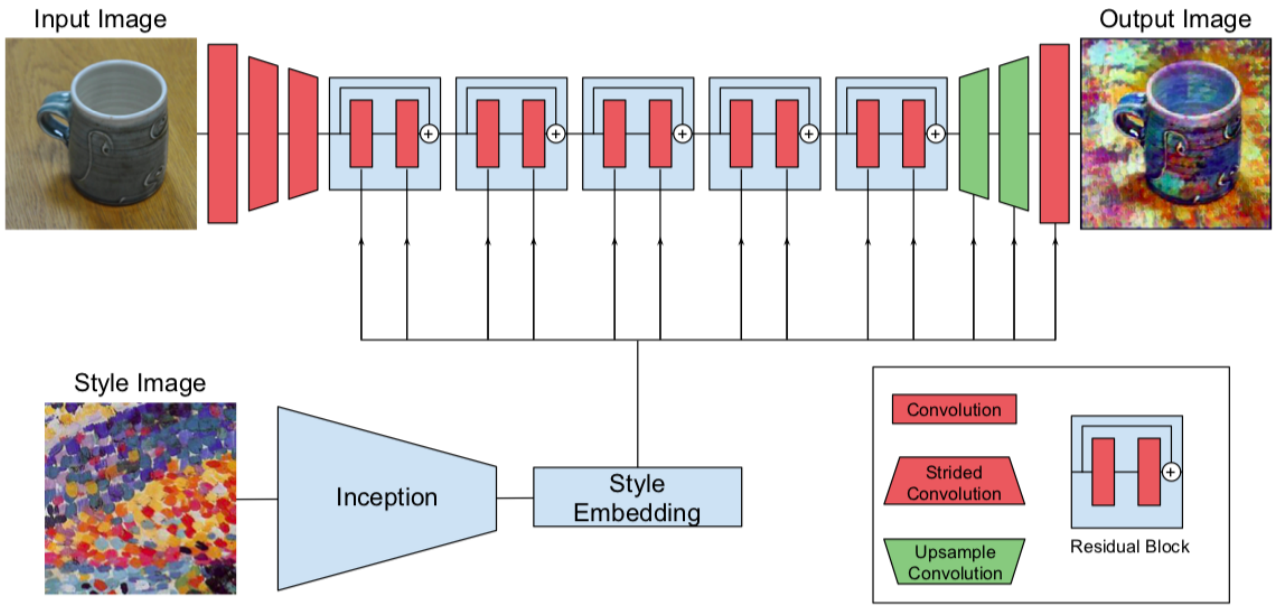
$$CIN(x) = \gamma^s \left(\frac{x - \mu_{nc}(x)}{\sigma_{nc}(x)} \right) + \beta^s$$

Сеть с кодировщиком стилей



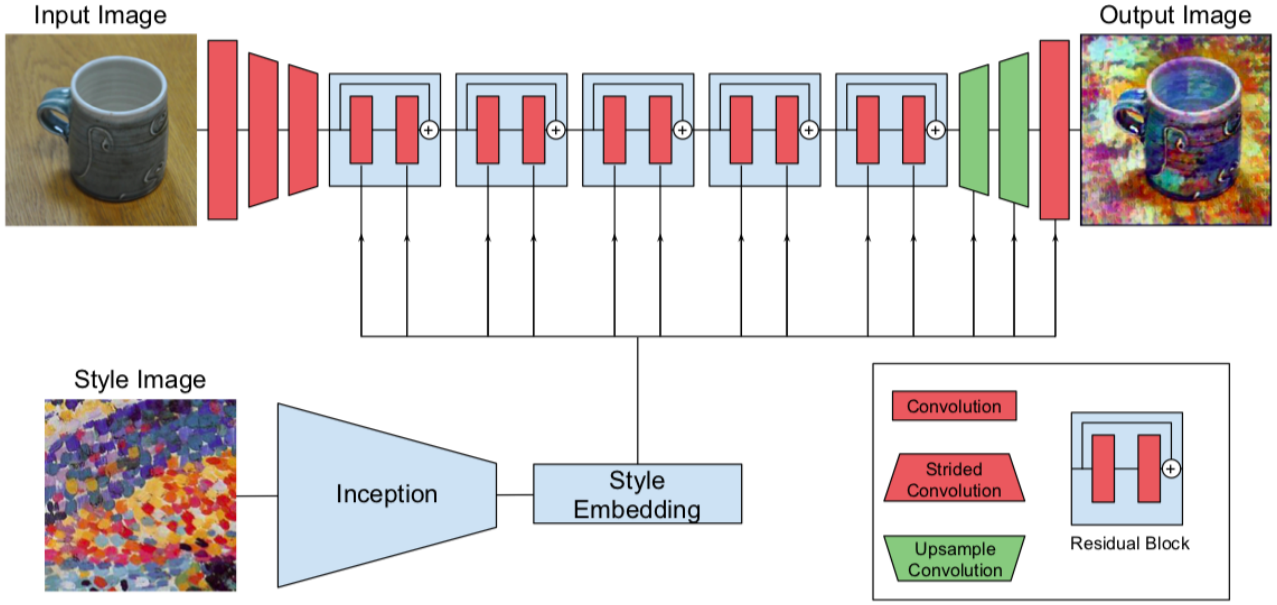
- Ghiasi et al. "Exploring the structure of a real-time, arbitrary neural artistic stylization network."(2017) [3]
- Кодировщик стилей
- Классические функции потерь
- Обобщение на произвольный стиль

Сеть с кодировщиком стилей



- Сеть-преобразователь $T(\cdot)$
  - Кодировщик-декодировщик
  - Остаточный слой: две свертки и две условной нормализации (CIN)
  - В CIN распределение контента подстраивается под распределение стиля

Сеть с кодировщиком стилей



- Кодировщик стиля $S(\cdot)$
  - Inception-v3 (mixed 6e)
  - Усреднение по каналам + Линейные слои($768 \to 100 \to 2758$)
  - Векторные представления стилей ($\mathbb{R}^{100}$ векторы) образуют богатое признаковое пространство

Обучение модели

- Три набора стилей: базовые, дополнительные, тестовые
- Обучение:
  - 20k итераций на базовых
  - 450k итераций на дополнительных с подмешиванием базовых с вероятностью 0.1
- Регуляризация: слой Dropout между Inception и слоем векторного представления стилей

Algorithm 1: Обучение модели на новых стилях.

init model $NN(\cdot)$;
init datasets $\mathcal{C}, \mathcal{S}_{base}, \mathcal{S}_{new}$;
init γ;
for $(I_c, I_s) \leftarrow (\mathcal{C}, \mathcal{S}_{new})$ **do**
  $\xi =$ sample from $U[0, 1]$;
  if $\xi > \gamma$ **then**
    $I_s =$ sample from $\mathcal{S}_{base}$;
  end
  $I_r = NN(I_c, I_s)$;
  compute loss, gradients and perform optimizer step;
end

Результат для базовых стилей

Content | Style | Baseline [2] | Ours

Рис.: Сравнение результатов стилизации на базовых стилях с бейзлайном.

Результат для базовых стилей

Content | Style | Baseline [2] | Ours

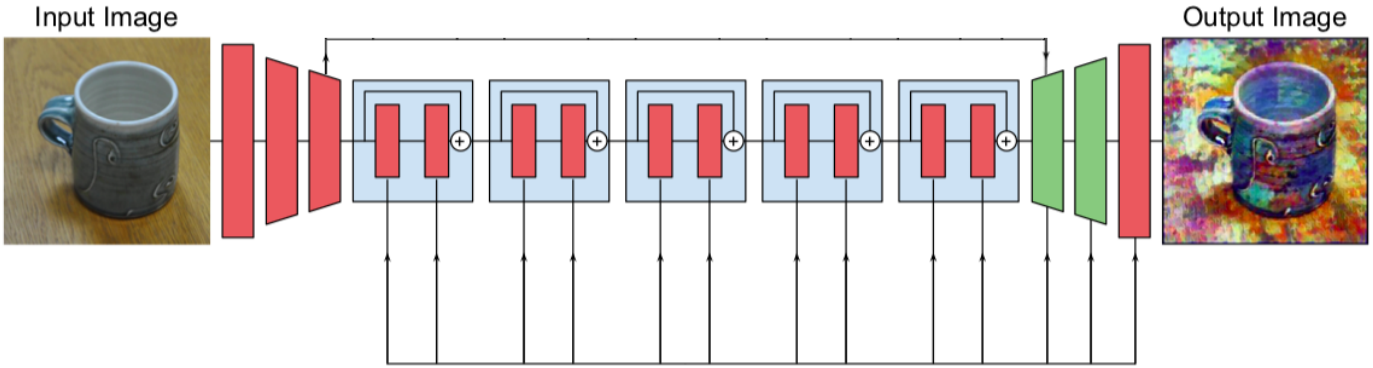
Рис.: Сравнение результатов стилизации на базовых стилях с бейзлайном.

Результат для тестовых стилей

Content | Style | Baseline [2] | Ours

Рис.: Сравнение результатов стилизации на тестовых стилях с бейзлайном.

Контроль силы стилизации модели



- Модификация: добавление связей из кодировщика в декодировщик
- $\alpha \in [0, 1]$ - параметр силы стилизации
  - Итоговое векторное представление стиля получается как линейная интерполяция:

$$\hat{s} = \alpha S(I_s) + (1 - \alpha) S(I_c)$$

  - Интерполяция связей:

$$x_{dec} = \alpha x_{res} + (1 - \alpha) x_{enc}$$

Контроль силы стилизации модели

Content | Style | $\alpha = 0.0$ | $\alpha = 0.3$ | $\alpha = 0.6$ | $\alpha = 1.0$

Рис.: Сравнение результатов работы метода при разном значении параметра силы стилизации.

Контроль силы стилизации модели с постпроцессингом

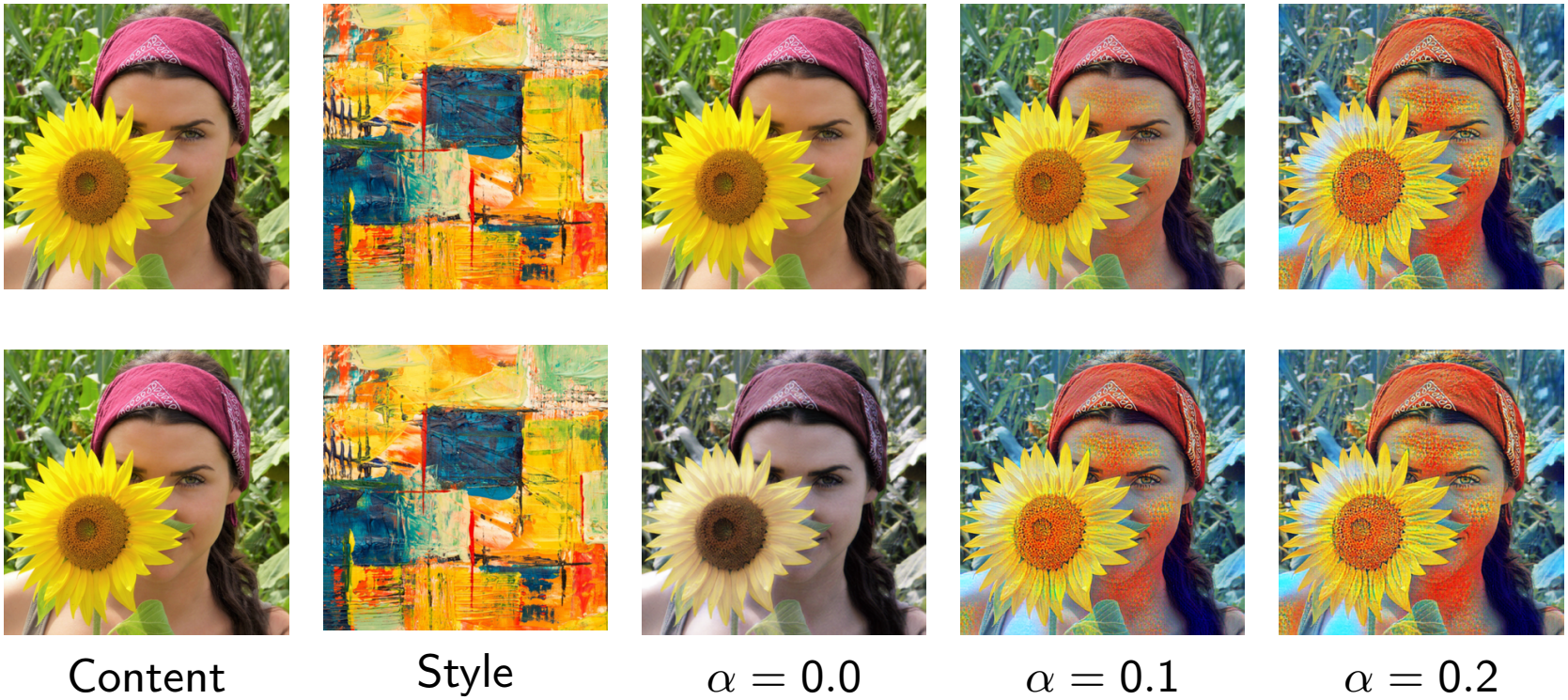


Рис.: Сравнение результатов работы метода с постпроцессингом (вверху) и без (внизу).

Постпроцессинг:

$$I_r = \frac{1}{\theta}\left(\alpha T(I_c) + (\theta - \alpha) I_c\right), \ \alpha < \theta$$

Контроль силы стилизации модели с постпроцессингом

Content | Style | $\alpha = 0.0$ | $\alpha = 0.1$ | $\alpha = 0.2$

Рис.: Сравнение результатов работы метода с постпроцессингом (вверху) и без (внизу).

Постпроцессинг:

$$I_r = \frac{1}{\theta}\left(\alpha T(I_c) + (\theta - \alpha) I_c\right), \ \alpha < \theta$$

Плавное сохранение цветов контента

- Задача: сохранение цветовой гаммы контента
- Методы
  - Стилизация только по каналу яркости в HCL представлении
  - Перекраска изображений: приведение гистограмм
- Второй способ позволяет осуществить плавность перекраски

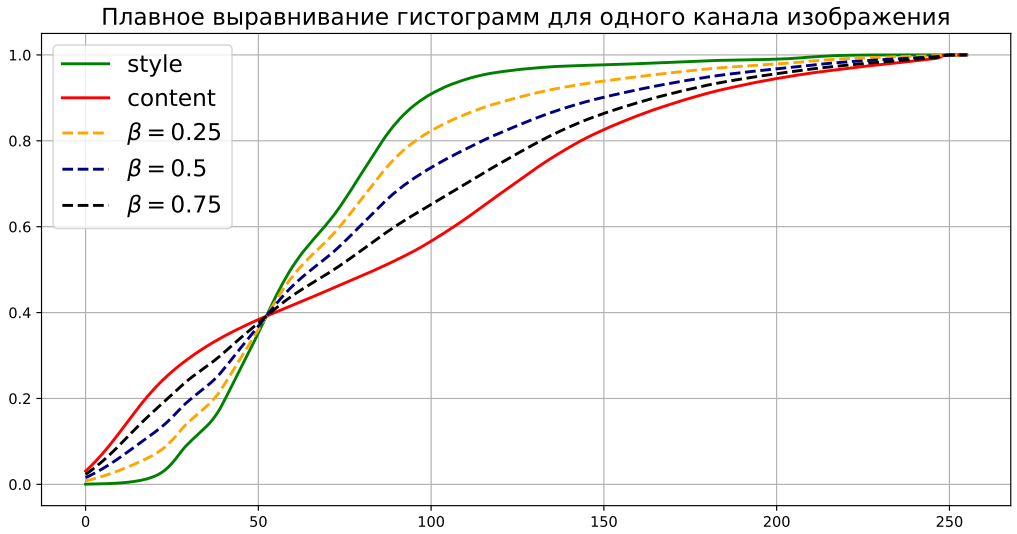


Рис.: Визуализация плавного приведения распределения стиля к распределению контента. Варьируя параметр β можно контролировать силу перекраски.

Контроль силы перекраски

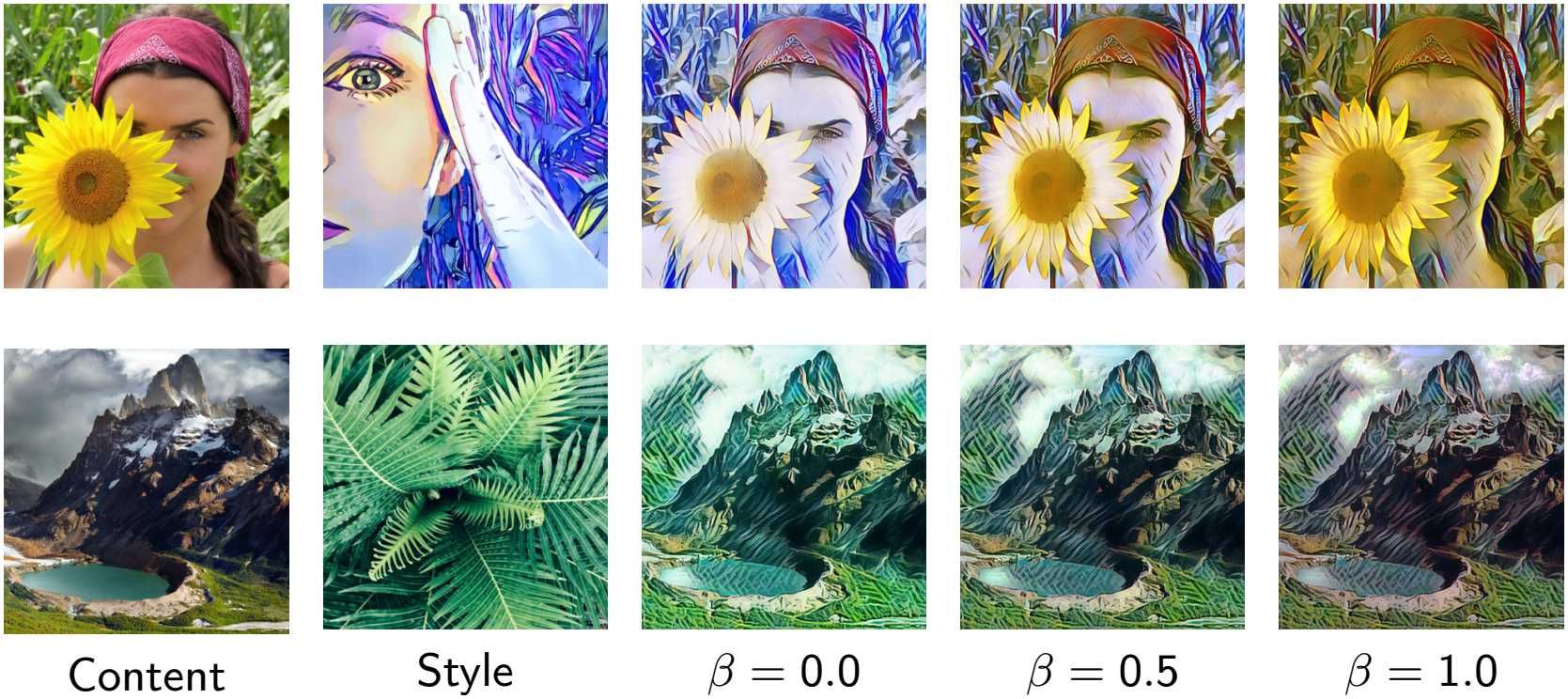


Рис.: Сравнение результатов работы метода при разном значении параметра силы перекраски.

Постпроцессинг: результат стилизации перекрашивается в цвета контента

Дополнительные ручки контроля стилизации

- Кластеризуем пространство векторных представлений для большой выборки стилей
- Каждый кластер соответствует некоторому стилевому паттерну
- Задавая веса при центроидах (или сэпмлах) каждого кластера и вычисляя их линейную интерполяцию, можно контролировать перенос стилевых паттернов

$$\hat{s} = \frac{\alpha}{\sum_{i=0}^{n} w_i} \left(S(I_s) w_0 + \sum_{i=1}^{n} s_i w_i \right) + (1 - \alpha) S(I_c)$$

Кластеризация

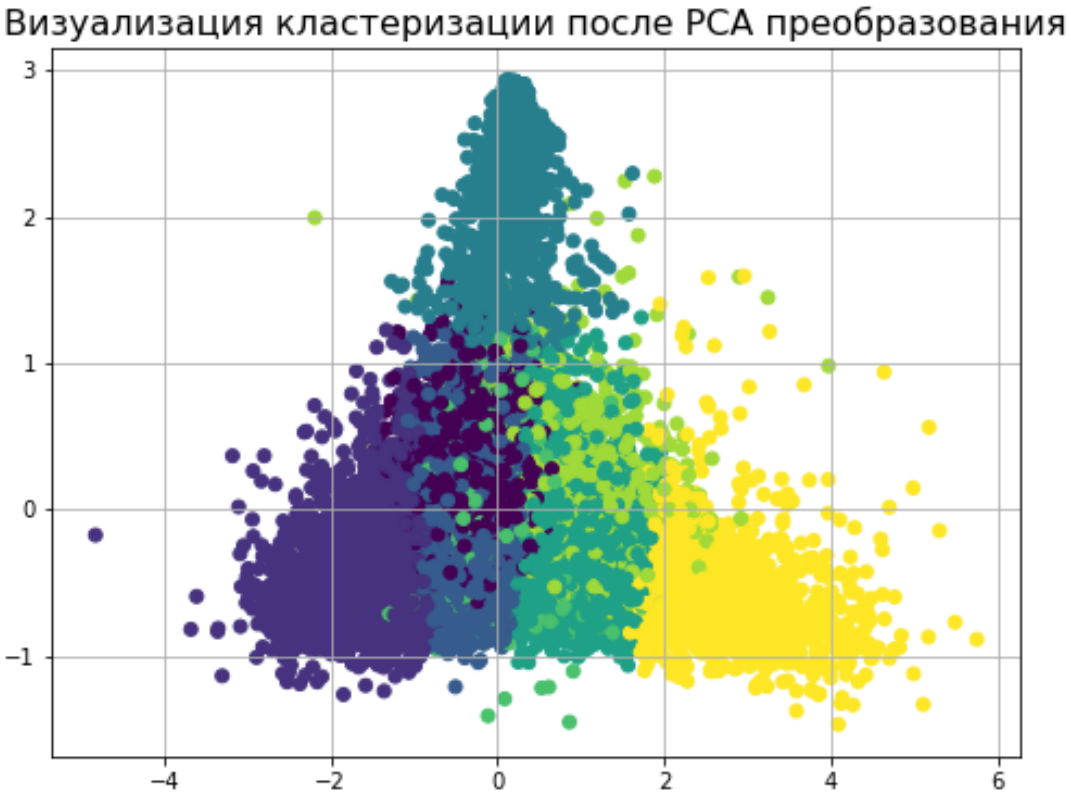


Рис.: Визуализация кластеризации в двумерном пространстве. Большинство кластеров имеют выпуклую форму. Каждый кластер соответствует некоторому стилевому паттерну характерному для объектов этого кластера.

Центроиды кластеров

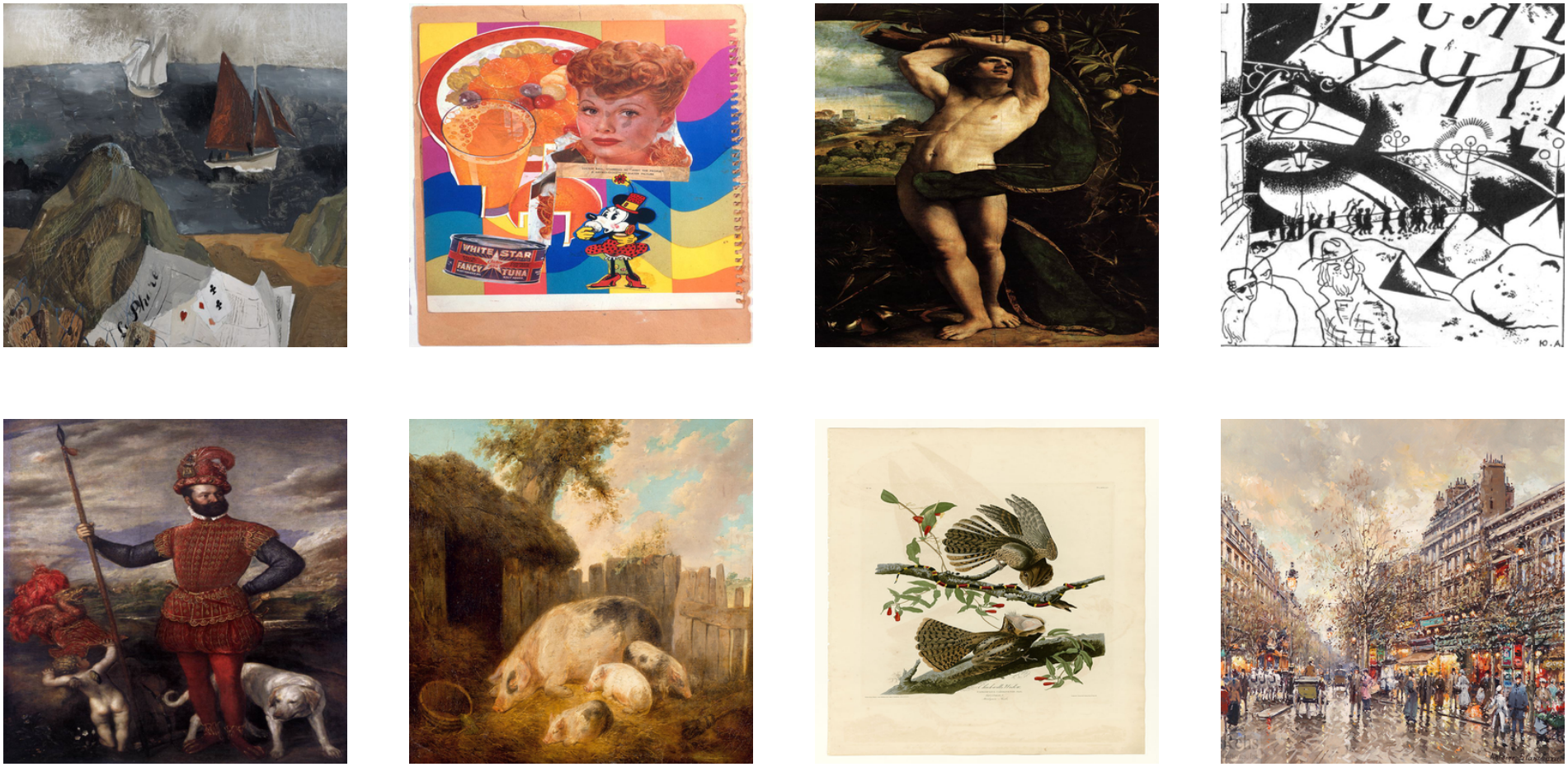


Рис.: Центроиды полученные после кластеризации стилей в пространстве векторных представлений.

Контроль переноса стилевых паттернов

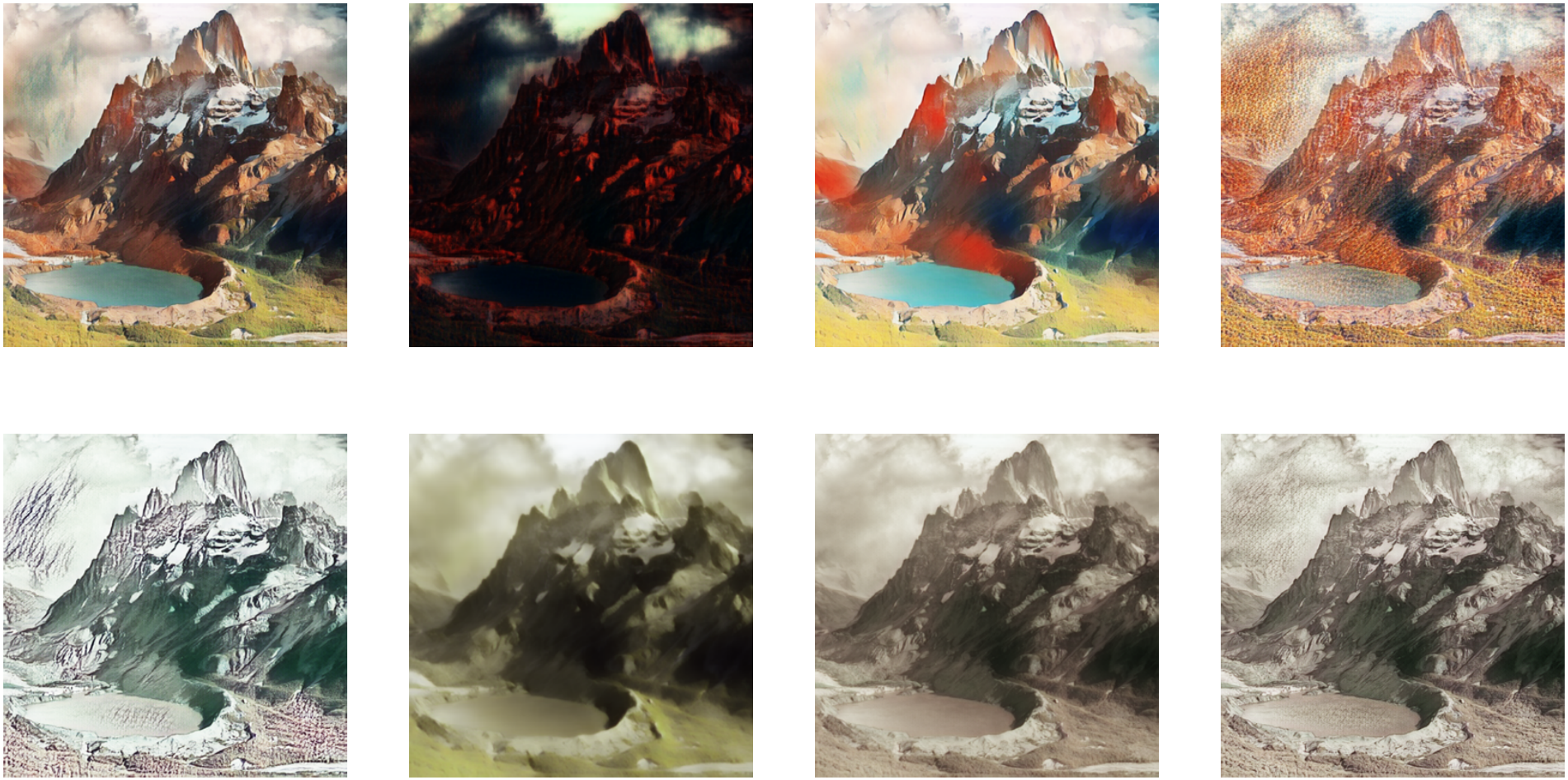


Рис.: Различные стилизации со случайными весами у центроид кластеров.

Ручки контроля стилизации

- Полный набор ручек контроля стилизации:
  - α - сила стилизации
  - β - сила перекраски
  - ω - набор весов для сэмплов или центроидов каждого кластера

Положения, выносимые на защиту

В рамках данной работы выполнено:

- Реализован метод с добавлением регуляризации и двухэтапной схемой обучения, работающий с произвольным стилем, обученный на втрое меньшем датасете относительно экспериментов авторов статьи
- Реализован метод контроля силы стилизации, сохраняющий изображение контента при малых значениях параметра
- Реализован плавный метод перекраски, позволяющий сохранять цветовую гамму контента
- Реализованы дополнительные параметры контроля стилизации: веса при векторном представлении центроид кластеров пространства стилей